Chapter 17

Data collection and representation

Contents:

OPENING PROBLEM

Tony and Carl play for the same basketball team. Due to an injury at practice Tony played only half of the season. The points scored by the players in each match were:
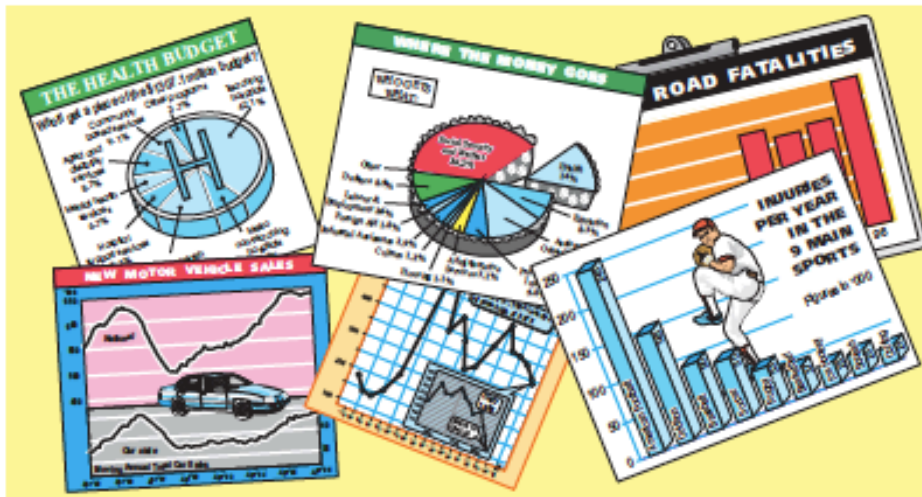
Tony: 17 21 15 8 18 12 27 15 22 31 28 8

Carl: 19 19 13 10 15 15 24 18 26 27 23 13
        20 24 18 26 19 25 8 36 21 23 26 19

Which player's performance was better?

Things to think about:

- Would it be fair to simply total the points each player scored for the season?
- How could we display the data in the meaningful way?
- What would be the 'best' way to solve the problem?

Statistics is about collection, organisation, display, analysis and interpretation of data.

Many groups such as schools, businesses and government departments collect information. The information is used to determine whether changes are needed, or whether changes that have been made have been successful.

Governments sometimes conduct a **census** in which they gather data from all of the nation's population. This information is used to help make decisions which will affect us in the future. For example, the government must consider how much money needs to be provided for health care in the years ahead because the number of elderly people is increasing.

Results of the collection and interpretation of data are displayed using **graphs**, **tables** and **diagrams**.

A   SAMPLES AND POPULATIONS

These are important words used in statistics:

Population: The whole group of objects or people from whom we are collecting data.

Sample: A group chosen to take part in a survey or to be measured or tested.

Random sample: A sample selected so that any person or object has as much chance as any other of being selected.

Inference: A conclusion you make based on your survey or investigation.

For example, suppose we conduct a survey on how much chocolate students at your school eat. The *population* is the students at your school.

A *sample* is chosen by selecting 10 students at random from the school roll. An *inference* might be that most students eat chocolate at least once a week.

CENSUS OR SURVEY

When the government carries out a **census** it requires everyone in the **population** to take part. This process is very expensive and takes a lot of time.

Instead, the government may conduct a **survey** of a **sample** of the population. It is important that the results of a survey are typical of the whole population. To ensure this, the sample must be randomly chosen, and as large as is practical.

DISCUSSION

1 Discuss why:

   a clothing manufacturers would like to know the body measurements of people in different age groups

   b the manager of your school canteen would be interested in the types and quantities of food you eat

   c your school keeps records of what is bought by the school population throughout the year

   d meteorologists are interested in temperature, rainfall, and atmospheric pressure measurements throughout the country and throughout the world.

2 For each of the situations listed in question **1**, discuss how the information could be collected.

3 Discuss how you would gather data in each of the following situations:

   a You wish to manufacture shoes and want to know how many of each size to make.

b As a private citizen you wish to make a case for traffic lights near the local school.

c You own a lawn mowing business and want to expand your business to a new area.

d You are an employer and you need to choose one person from 50 applicants.

4 Organisations and marketing researchers have many clever ways of gathering information by tempting us with offers. Discuss some of the ways in which information is collected from you. Collect samples from newspapers, magazines, packaging, and letterbox deliveries which invite you to provide data.

EXERCISE 17A

1 Suggest how to select a random sample of:

    a 400 adults   **b** bottles of soft drink at a factory

    c 30 students at a school   **d** words from the English language

    Are there any advantages or disadvantages in the methods you have suggested?

2 How would you randomly select:

    a one ticket out of 5 tickets

    b one of the letters A or B

    c one of the numbers 1, 2, 3, 4, 5 or 6

    d a card from a pack of 52 playing cards?

Example 1　　　　　　　　　　　　**◀》 Self Tutor**

From a school of 400 students, a random sample of 60 students was selected. 13 were found to have blue eyes.

    a How many students are in the population?

    b How many students are in the sample?

    c What fraction of the sample has blue eyes?

    d Estimate how many in the population have blue eyes.

You must know the difference between a population and a sample.

    a There are 400 students in the population.

    b There are 60 students in the sample.

    c 13 out of 60 students in the sample have blue eyes

        so the fraction of the sample with blue eyes is $\frac{13}{60}$.

    d $\frac{13}{60}$ of 400 $\{\frac{13}{60}$ of the population have blue eyes$\}$

      $= \frac{13}{60} \times 400$

      ≈ 87

    Calculator: 13 $\boxed{\div}$ 60 $\boxed{\times}$ 400 $\boxed{=}$

    So, approximately 87 students in the school have blue eyes.

3 From a colony of 10 000 ants, 300 are collected and examined for red eye colour. 36 were found to have red eyes.

    a How many ants form the population?

    b How large was the sample?

    c What percentage of the sample had red eyes?

    d Estimate the total number of red-eyed ants.

4 50 people were randomly selected from the 750 who attended the opening night of a new play. Of the 50 people, 33 said that they liked the play.

    a What was the population size of people attending the play?

    b How large was the sample?

    c What percentage of the sample did *not* like the play?

    d Estimate the total number of people who did not like the play.

B CATEGORICAL DATA

Categorical data is data which can be placed in categories.

For example, suppose we stand at a street intersection and record the colour of each car going past.

We use the code R = red, B = blue, G = green, W = white, O = other colours to help us record the data efficiently.

The following results were observed in a sample of 50 cars:

    BGWWR   OGWRW   OOBBG   OGRWR   WWWGB

    BBGGW   WWWOG   WOBWW   RWWRB   OOBWR

Having collected our categorical data, we first **organise** it in groups. We can do this using either
- a **dot plot** or
- a **tally and frequency table**

Organisation of the data helps us to identify its features. For example:

The **mode** is the most frequently occurring category.
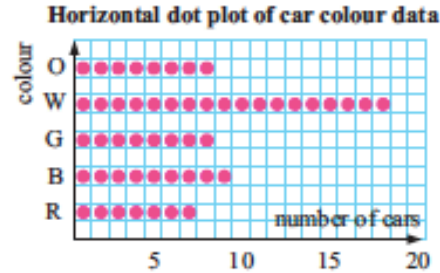
DOT PLOTS

A **dot plot** is a graph which displays data, where each dot represents one data value.

Dot plots are often used to record data initially and may be **horizontal** or **vertical**.

A dot plot for the car colour data is shown alongside:

Check that there is one dot for each car recorded in the data.

The mode is 'white' as W is the most frequently occurring category.

Horizontal dot plot of car colour data
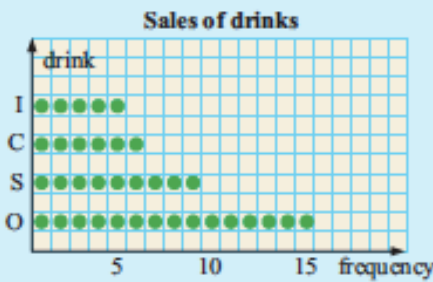

Example 2   ◀) **Self Tutor**

At recess time the sales of drinks were recorded over a three minute period.
O = 100 plus,   S = soy milk,   C = cola,   I = iced tea.

The data was: OSSCI   OCISO   IOCSO   OOOSC   SOCOS   SOOCO   OIOIS

a Draw a dot plot of the data.   **b** What is the mode?

a

Sales of drinks

b The mode is 100 plus.

EXERCISE 17B.1

1 The dot plot shows shoe sizes for students in grade 6.

  a How many students are in grade 6 at this school?

  b How many have shoe size 9 or more?

  c What percentage have shoe sizes 8 or more?

Shoe sizes

2 The dot plot shows the numbers of students playing various instruments in the school orchestra.

  a How many play stringed instruments?

  b How many students are in the orchestra?

  c Find the mode of the data.

School orchestra
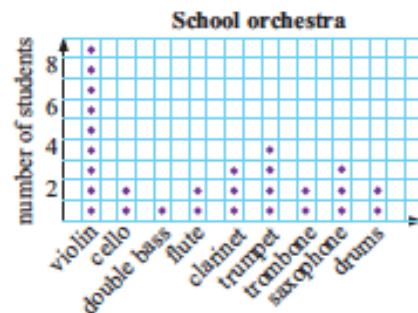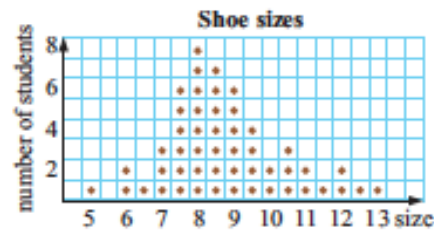
3 A class of students at a school in England were asked which summer sport they wanted to play. The choices were: T = tennis,   S = swimming,   C = cricket,   B = basketball and   F = football.

The data was: FFCTC   CSFST   TTBFS   FFCSF   TFTBC

   a Draw a horizontal dot plot of the data.

   b Find the mode of the data.

4 Students voted the most popular attractions at the local show to be the side shows (S), the farm animals (F), the ring events (R), the dogs and cats (D), and the wood chopping (W). The students in a class were then asked to name their favourite.

The data was: SRWSS   WFDDS   RRFWS   RSRWS   SRRRF

   a Draw a vertical dot plot of the data.

   b Find the mode of the data.

TALLY AND FREQUENCY TABLES

If there is a lot of data, a tally and frequency table is a useful way to collect the information.

The **tally** is used to count the data in each category. The **frequency** summarises the tally, giving the total number of each category.

Such a table is also called a **frequency distribution table** or simply a **frequency table**.

For the car colour data the frequency table is:

Colour	Tally	Frequency
Other	卌 |||	8
White	卌 卌 卌 |||	18
Green	卌 |||	8
Blue	卌 ||||	9
Red	卌 ||	7

The tally uses strokes to record each result.

> The **frequency** of a category is the number of items in that category.

Example 3   ◀》 **Self Tutor**

The data below records how students in a class travel to school on a particular day.

W = walk,   Bi = bicycle,   Bu = bus,   C = car,   T = train

The data is:

W Bi Bu T C   Bi C W Bi Bu   Bi C C Bi Bu   W Bu Bu T C
Bi Bi Bu T C   C Bi C C C   W W Bu T C

   a Draw a frequency table to organise the data.

   b Find the mode of the data.

a

Method of Travel	Tally	Frequency
Walk	Ḩ̶Ħ	5
Bicycle	Ḩ̶Ħ |||	8
Bus	Ḩ̶Ħ ||	7
Car	Ḩ̶Ħ Ḩ̶Ħ		11
Train	||||	4
	Total	35

b The mode is 'car' as this category occurs most frequently.

EXERCISE 17B.2

1 The results of a survey of eye colour in a class of 28 year 6 students were:

Br Bl Gn Bl Gn   Br Br Bl Gn Gr   Br Gr Br Br Bl   Br Bl Br Gr Gn
Br Bl Br Gn Gr   Br Bl Gn

where   Br = brown,   Bl = blue,   Gn = green,   and   Gr = grey.

   a Complete a frequency table for the data.

   b Find the mode of the data.

2 Students in a science class obtained the following levels of achievement:

D C C A A C C D C B C C C D B C C C C E B A C C B C B C

   a Complete a frequency distribution table for the data above.

   b Use your table to find the:

   i number of students who obtained a C

   ii fraction of students who obtained a B.

   c What is the mode of the data?

3 Tourists staying in a city hotel were surveyed to find out what they thought about the service by the hotel staff. They were asked to choose   E = excellent, G = good, S = satisfactory, or U = unsatisfactory. The results were:

EGGSE   USSGG   SGUGG   ESGUG   SSEGG

   a Complete a frequency table for the data.

   b What is the mode of the data?

   c Suggest a reason why this survey would be carried out.

C   GRAPHS OF CATEGORICAL DATA

Categorical data is often displayed using **column graphs** and **pie charts**.

USING HAESE & HARRIS SOFTWARE

Click on the icon to load a statistical package which can be used to draw a variety of statistical graphs.

Change to a different graph by clicking on a different icon. See how easy it is to change the labels on the axes and the title of the graph.

STATISTICS
PACKAGE

Type into the correct cells the information given on car colour.

Colour	Frequency
white	38
red	27
blue	19
green	18
other	11

Print off graphs of the data including a pie chart, a column graph, and a strip graph.

Use the software or a spreadsheet to reproduce some of the statistical graphs in the remaining part of this chapter. You can also use this software in any statistical project you may be required to do.

COLUMN GRAPHS

Column graphs consist of rectangular columns of equal width. The height of each column represents the the frequency of the category.

DEMO

Example 4
Self Tutor

The graph given shows the types of drink purchased by students at recess time.

Recess time drinks

a What is the least popular drink?

b What is the mode of the data?

c How many students drink orange juice?

d What percentage of students drink chocolate milk?

a Iced coffee   {shortest column}

b 'Soft drink' is the mode.

c 27 students drink orange juice.

d The total number of students purchasing drinks $= 27 + 35 + 18 + 10 = 90$

So, the percentage of students drinking chocolate milk is $\frac{18}{90} \times 100\% = 20\%$

EXERCISE 17C.1

1 The results of a survey of eye colour in a class of 28 year 6 students were:

Eye colour	Brown	Blue	Green	Grey
Frequency	11	7	6	4

a Illustrate these results using a hand drawn column graph.

b What is the most frequently occurring eye colour?

c What percentage of the students have blue eyes?

2
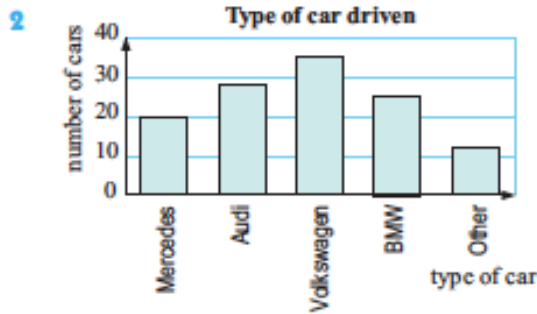
Type of car driven

The column graph shows the type of vehicle driven by 120 randomly selected people in Munich.

a Use the graph to estimate the frequency of each type of car.

b Which make of car is the most popular?

c What percentage of the surveyed people drive an Audi?

3 Yearly profit and loss figures for a business can be easily illustrated on a column graph. For the example given:

a in what years was a profit made

b what happened in 2006

c what was the overall profit or loss over the 6-year period?

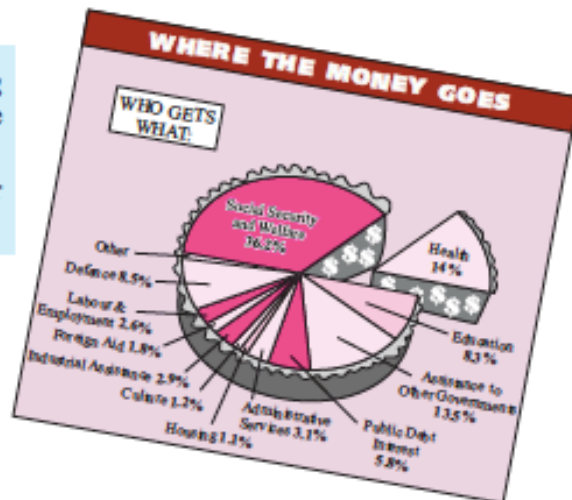**Profits for XNP Nominees**
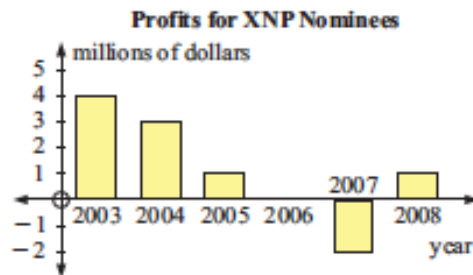
PIE CHARTS

A **pie chart** is a useful way of displaying how a quantity is divided up. A full circle represents the whole quantity.

We divide the circle into **sectors** or wedges to show each type or category.

For example, the pie chart alongside shows how the budget of a country is distributed.


Fruit	Frequency
Orange	13
Apple	21
Banana	10
Pineapple	7
Pear	9
Total	60

The table opposite shows the results when a class of year 8 students were asked 'What is your favourite fruit?'.
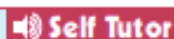
There are 60 people in the sample, so each person is entitled to $\frac{1}{60}$th of the pie chart. $\frac{1}{60}$th of $360°$ is $6°$, so we can calculate the sector angles on the pie chart:

$13 \times 6° = 78°$ for the orange sector

$21 \times 6° = 126°$ for the apple sector

$10 \times 6° = 60°$ for the banana sector

$7 \times 6° = 42°$ for the pineapple sector

$9 \times 6° = 54°$ for the pear sector.
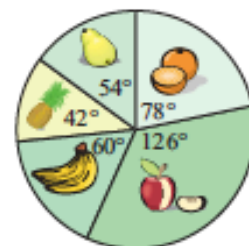
The completed pie chart is shown alongside.

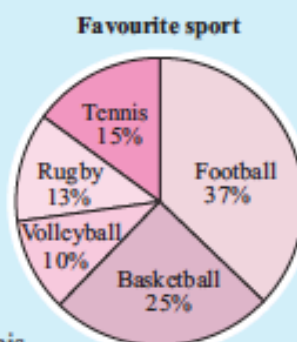**Example 5**   ◀ᴐ **Self Tutor**

The pie chart shows the results of a survey of 120 year 7 students. The students were asked the question: *"What is your favourite sport?"*

Use the chart to determine:

 a the most popular sport

 b the least popular sport

 c the number of students whose favourite sport is basketball

 d the number of students whose favourite sport is tennis.

Favourite sport

 a The largest sector angle indicates the most popular sport, and this is football.

 b Volleyball is the least popular sport as it has the smallest sector angle.

 c 25% of students said basketball is their favourite sport.
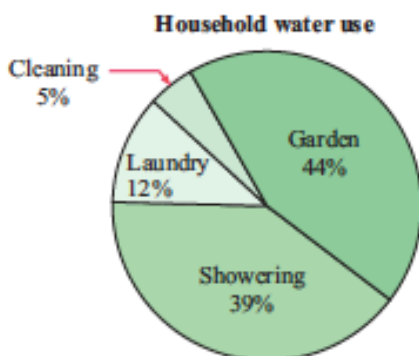 So, the number of students whose favourite sport is basketball is

  25% of 120

  $= \frac{25}{100} \times 120 = 30$ *Calculator:* 25 ÷ 100 × 120 =

 d The number of students whose favourite sport is tennis is

  15% of 120

  $= \frac{15}{100} \times 120 = 18$ *Calculator:* 15 ÷ 100 × 120 =

EXERCISE 17C.2

1

Household water use

The pie chart alongside illustrates the proportion of water required for various household uses.

 a For what purpose is the most water used?

 b For what purpose is the least amount of water used?

 c If the household used 400 kilolitres of water during a particular period, estimate the quantity of water used for:
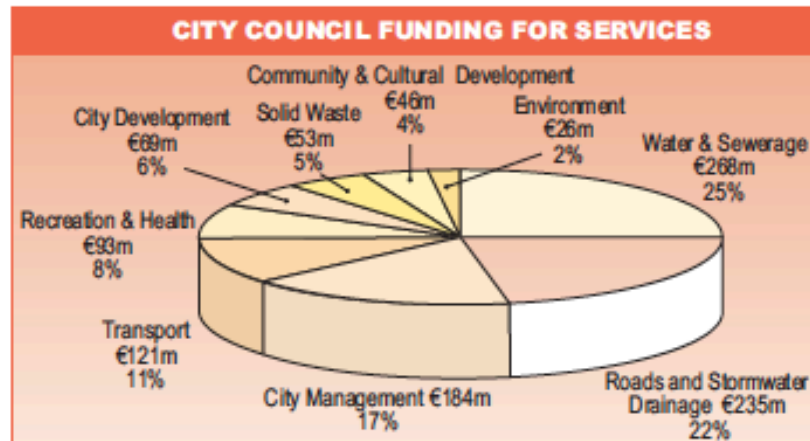
  i showering   **ii** cleaning.

2 The pie chart alongside shows the percentages of women in France who wear certain sizes of clothing.

 a Find what size is most commonly worn.

 b A group of 200 women attends a fashion parade. Estimate how many would wear size 10 clothing.

Clothing sizes
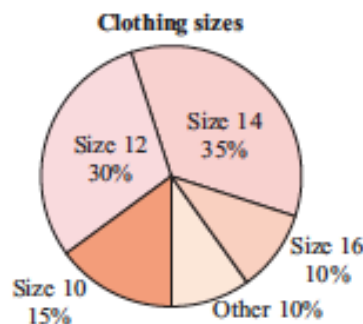
3 This pie chart shows both the percentages and the actual amounts a council spent in various areas.

CITY COUNCIL FUNDING FOR SERVICES

Community & Cultural Development €46m 4%

City Development €69m 6%

Solid Waste €53m 5%

Environment €26m 2%

Water & Sewerage €268m 25%

Recreation & Health €93m 8%

Transport €121m 11%

City Management €184m 17%

Roads and Stormwater Drainage €235m 22%

a Briefly describe what the graph is about.

b Comment on the usefulness of having both percentages and amounts shown.

c What percentage of total funding is spent on:
  i Recreation and Health   **ii** Community and Cultural Development?

d How much money is spent on:
  i Environment   **ii** City Development?

e On what service is the largest amount spent?

f How much is spent in total?

D   NUMERICAL DATA

Numerical data is data which is in number form.

Numerical data can be *organised* using a **stem-and-leaf plot** or a **tally and frequency table**. Numerical data is usually represented graphically by a **column graph**.

STEM-AND-LEAF PLOTS

A **stem-and-leaf plot** can be used to write a set of data in order.

For example, the weights (in kg) of army recruits are:

101, 91, 83, 84, 72, 93, 67, 85, 79, 87, 78, 89, 68,
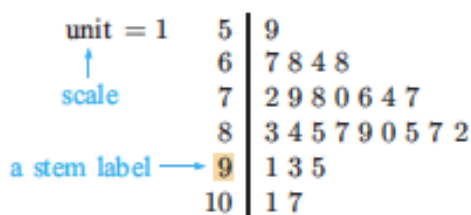80, 107, 70, 85, 64, 95, 76, 87, 74, 68, 59, 82, 77

For each data value, the units digit is used as the **leaf**, and the digits before it determine the **stem** on which the leaf is placed.

So the stem labels are 5, 6, 7, 8, 9, and 10, and they are written under one another in ascending order.
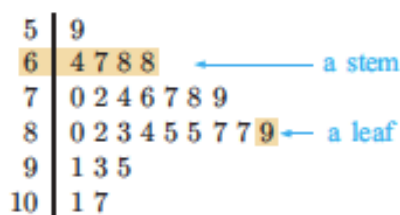
We now look at each data value in turn. The first data value is 101. Its stem label is 10 and its leaf is 1. We record 1 to the right of the stem label 10. The next data value is 91. Its stem label is 9 and its leaf is 1. We record 1 to the right of the stem label 9. Using this method we record all the data in an unordered stem-and-leaf plot.

Unordered stem-and-leaf display of weight data

```
unit = 1    5 | 9
   ↑        6 | 7 8 4 8
 scale      7 | 2 9 8 0 6 4 7
            8 | 3 4 5 7 9 0 5 7 2
a stem label → 9 | 1 3 5
           10 | 1 7
```

Ordered stem-and-leaf display of weight data

```
            5 | 9
            6 | 4 7 8 8    ←——— a stem
            7 | 0 2 4 6 7 8 9
            8 | 0 2 3 4 5 5 7 7 9 ←— a leaf
            9 | 1 3 5
           10 | 1 7
```

The leaves on each stem are now written in ascending order. So for the stem 6 | 7 8 4 8 we write 6 | 4 7 8 8.   This gives an ordered stem-and-leaf plot.

Notice that:

- 6 | 4 7 8 8   represents the four scores 64, 67, 68 and 68.
- The leaves are placed in ascending order.
- The scale (unit = 1) tells us the place value of each leaf. If the scale was 'unit = 0.1' then 6 | 4 7 8 8 would represent 6.4, 6.7, 6.8, 6.8
- Rotating the diagram, we see the shape of a column graph.

```
 5 | 9
 6 | 4788
 7 | 0246789
 8 | 023455779
 9 | 135
10 | 17
```
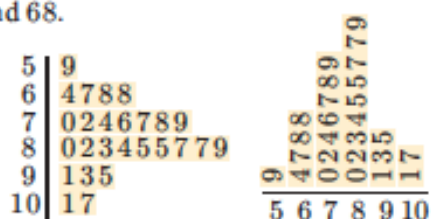
Example 6   🔊 Self Tutor
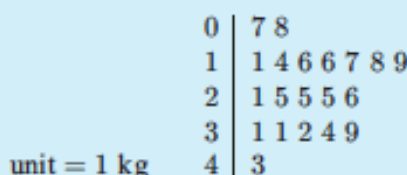
A fisherman recorded the total weight of all schnapper he caught each day. Construct a stem-and-leaf plot for the data shown below (in kg):

11, 16, 07, 25, 39, 26, 14, 17, 18, 31
31, 25, 43, 32, 25, 19, 16, 08, 34, 21

Stem-and-leaf display for schnapper catch

```
              0 | 7 8
              1 | 1 4 6 6 7 8 9
              2 | 1 5 5 5 6
              3 | 1 1 2 4 9
unit = 1 kg   4 | 3
```

EXERCISE 17D.1

1

The weights of 24 football players were recorded to the nearest kg as follows:

```
72   63   90   70   67   71   89   64   93   86
66   78   75   89   80   91   81   72   87   72
86   84   84   87
```

Construct a stem-and-leaf plot to display this data.

2 The weights of 30 fifteen week old piglets were recorded to the nearest kg as follows:

18 20 30 30 25 19 30 34 28 36 32 33 38 13 37
29 43 50 20 44 23 27 27 47 37 17 38 51 29 39

Construct a stem-and-leaf plot to display this data.

3 The time in hours taken by a farmer to plough, fertilise, and seed each of his paddocks is given below:

7 24 9 12 41 30 36
28 18 27 32 24 13 25

Construct a stem-and-leaf plot to display this data.

4 The time (in hours) taken by farmers to travel to their nearest town centre is given below:

1.0 2.4 0.9 1.2 3.6 3.0 0.7
0.8 1.8 2.7 0.2 2.4 1.3 0.5

Construct a stem-and-leaf plot to display the data, stating the scale used.

WORKING WITH NUMERICAL DATA

Example 7   ◀️ Self Tutor

An exceptional hockey player scores the following number of goals over a 25 match period: 4 3 6 1 5 8 4 2 2 4 6 0 5 1 9 3 7 2 6 6 8 3 6 2 10

a Organise the data in a tally and frequency table.
b Graph the data on a column graph.
c On how many occasions did the player score 5 or more goals in a match?
d On what percentage of occasions did the player score 4 or more goals in a match?

a

| Goals	Tally	Frequency
0			1
1				2
2						4
3					3
4					3
5				2
6	卌	5
7			1
8				2
9			1
10			1

b

Hockey goal data
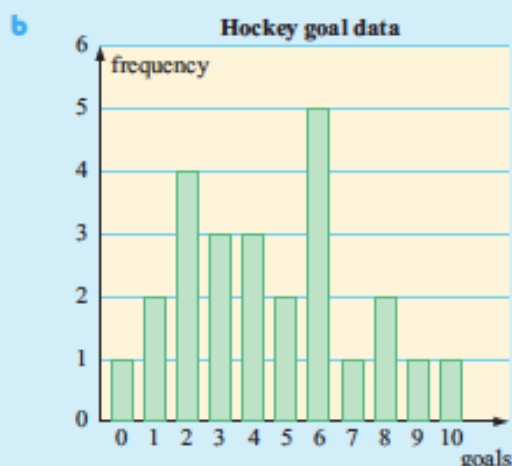
> **c** $2 + 5 + 1 + 2 + 1 + 1 = 12$ times
>
> **d** He scored 4 or more goals on 15 occasions.
> So, the percentage $= \dfrac{15}{25} \times 100\%$
> $= \tfrac{3}{5} \times 100\%$
> $= 60\%$

EXERCISE 17D.2

1 a Complete a frequency distribution table for the number of children in 30 families:
$$0, 4, 6, 2, 1, 3, 2, 4, 0, 2, 1, 2, 5, 0, 2, 3, 1, 4, 2, 1, 2, 4, 3, 3, 0, 4, 5, 2, 2, 4$$
 b Use your table to find the:
   i number of families with two children
   ii fraction of families with three children.

2 Following are the ages of children at a party:
$$12, 11, 17, 12, 14, 13, 11, 12, 15, 13, 12, 14, 11, 14, 12, 10, 12, 11, 13, 14$$
 a Organise the data in a tally and frequency table.
 b How many children attended the party?
 c How many of the children were aged 12 or 13?
 d What percentage were 13 or more years old?
 e Display the data on a column graph.

3 The given graph shows the number of points scored by a basketballer over a 60-match period.
 a What point score occurred most frequently?
 b On how many occasions were 10 or more points scored?
 c In what percentage of matches were fewer than 5 points scored?

Basketball point scores
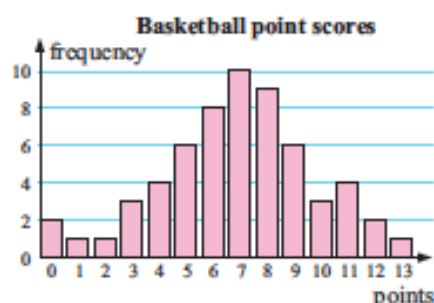
4 The numbers of goals kicked by a football player each match for the 2008 season were:
$$3 \quad 0 \quad 4 \quad 2 \quad 0 \quad 3 \quad 3 \quad 1 \quad 2 \quad 1 \quad 1 \quad 2 \quad 3 \quad 3 \quad 2 \quad 2 \quad 5 \quad 0 \quad 2 \quad 1 \quad 4 \quad 3$$
 a Complete a frequency table for the given data.
 b Use the table to find the number of games where the player kicked:
   i exactly 3 goals   **ii** at least 3 goals.

5 A record was kept of the number of goals scored by a goal shooter in netball games during the season. The results were:
$$10, 7, 8, 5, 8, 7, 10, 10, 6, 11, 5, 7, 7, 12, 7, 11, 6, 5, 8, 8, 7$$
 a Complete a frequency table for the data above.
 b Draw a column graph of the data.
 c Find the number of games in which the shooter scored:
   i exactly 8 goals   **ii** at least 8 goals.
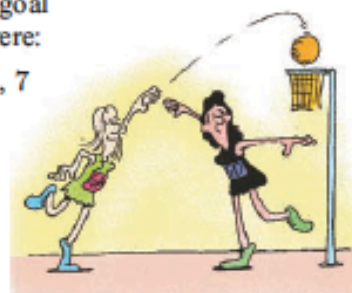
6 It is stated on match-boxes that the average contents is 50. When 40 boxes were sampled, the following numbers of matches were counted:

48, 51, 49, 50, 51, 52, 50, 48, 49, 51, 50, 53, 48, 49, 51, 50, 52, 49, 50, 52,
51, 48, 50, 49, 50, 51, 52, 50, 49, 48, 52, 50, 51, 49, 50, 50, 48, 53, 52, 49

  a Prepare a frequency table for this data.

  b How many boxes had exactly 50 matches?

  c How many boxes had 50 or more matches?

  d What fraction of boxes had less than 50 matches?

  e Do you think the manufacturer's claim is valid?

E | MEAN OR AVERAGE

The **mean** or **average** of a set of numbers is an important measure of their middle. We talk about averages all the time. For example:

- the average speed of a car
- average height and weight
- the average score for a test
- the average wage or income.

> The **mean** or **average** is the total of all scores divided by the number of scores.

For example, the mean of 2, 3, 3, 5, 6 and 11 is

$$\frac{2+3+3+5+6+11}{6} \quad \longleftarrow \quad \{\text{there are 6 scores}\}$$
$$= \frac{30}{6}$$
$$= 5$$

DISCUSSION   COMPARING DATA

Discuss how averages can be used to compare different sets of data. You may wish to consider these statements:

- In the last World Cup, Brazil scored an average of 2.3 goals per match. Germany scored an average of 1.5 goals per match.
- The X8 model travels 11.6 km per litre of fuel, whereas the Z3 travels 12.7 km per litre.
- In American Football, why is the average height and weight of the players important?

Example 8   ◀╢ **Self Tutor**

Find the mean of 7, 11, 15, 6, 11, 19, 23, 0 and 7.

$$\text{Mean} = \frac{7 + 11 + 15 + 6 + 11 + 19 + 23 + 0 + 7}{9}$$
$$= \frac{99}{9}$$
$$= 11$$

> The mean is a measure of the middle of a set of scores.

EXERCISE 17E

1 Find the mean of 1, 2, 3, 4, 5, 6 and 7.

2 Calculate the mean of the scores 7, 8, 0, 3, 0, 6, 0, 11 and 1.

3 The weights of a group of newborn ducklings are: 60 g, 65 g, 62 g, 71 g, 69 g, 69 g
Find the average birthweight of the ducklings.

4 In a skijumping competition, Lars jumps the following distances: 110 m, 112 m, 118 m, 103 m, 122 m
Calculate the average length of Lars' ski jumps.

5 In a basketballer's last 12 games of a season he scored 23, 18, 36, 29, 38, 44, 18, 52, 47, 20, 50, and 42 points. What was his mean point score over this period?

6 Baseballers Sean and Rick each throw a set of baseball pitches. The speeds of their pitches, in kilometres per hour, are:
Sean: 130, 135, 131, 119, 125   Rick: 132, 125, 138, 121, 129

   a Find the average speed of the pitches thrown by each baseballer.

   b Who has the fastest average pitching speed?

7 Compare the performance of two groups of students in the same mental arithmetic test out of 10 marks.
Group X: 7, 6, 6, 8, 6, 9, 7, 5, 4, 7   *Group Y:* 9, 6, 7, 6, 8, 10, 3, 9, 9, 8, 9

   a Calculate the mean of each group.

   b There are 10 students in *group X* and 11 in *group Y*. Because of unequal numbers in each group it is unfair to compare their means. True or false?

   c Which group performed better at the test?

8 The given data shows the goals scored by girls in the local netball association.

   a Find the mean number of goals for each goal shooter.

   b Which goal shooter has the best average performance?

Name	Goals	Games
Sally Brown	238	9
Jan Simmons	235	10
Jane Haren	228	9
Peta Piper	219	7
Lee Wong	207	8
Polly Lynch	199	7
Sam Crawley	197	6

ACTIVITY   **A POSSIBLE STATISTICAL EXPERIMENT**

In this activity you will grow wheat over a 21 day period in a controlled experiment. You will use 6 grains of wheat in each of 4 plots.

You will need: 4 saucers or coffee jar lids, cotton wool, 24 grains of wheat, measure, eye dropper, diluted liquid fertiliser.

What to do:

1   Layer the cotton wool three quarters of the way up each lid. Place 6 grains of wheat at equal distances apart in each lid.

2   Label the lids as plots 1, 2, 3 and 4. Saturate each plot with 15 mL of water.

3   Over a 3 week period, perform the following:

- In plot 1 squeeze 2 drops of water onto each grain of wheat every weekday.
- In plot 2 squeeze 2 drops of water onto each grain of wheat every Monday, Wednesday and Friday.
- In plot 3 squeeze 2 drops of water and 1 drop of diluted fertiliser onto each grain every Monday, Wednesday and Friday.
- In plot 4 squeeze 2 drops of water and 1 drop of diluted fertiliser onto each grain every weekday.

4   Place all the plots in the same safe, sheltered place with plenty of light.

5   Every Monday, Wednesday and Friday, record the mean height of any germinating seeds for each plot. Avoid handling any shoots. Make a table to summarise your results.

6   Use graphs and the language of statistics to comment on your results.

KEY WORDS USED IN THIS CHAPTER

- average
- dot plot
- inference
- numerical data
- random sample
- tally and frequency table
- bar graph
- frequency
- mean
- pie chart
- sample
- categorical data
- frequency table
- mode
- population
- stem-and-leaf plot

REVIEW SET 17A

1 The data below represents birth months in a year 7 class. January is represented by the number 1, February by the number 2, and so on up to December which is 12. Boys are shown in black and girls in blue.

<p align="center">6 7 3 9 5 5 9 12 10 4 1 12 6 3 5</p>
<p align="center">7 7 4 10 3 7 1 9 5 9 4 8 7 11 4</p>

  a Prepare a tally and frequency table to show this data.

  b Answer the following questions:

    i How many students were in the class?

    ii How many girls were in the class?

    iii What fraction of the class was born in April?

    iv What percentage of the class was born in March?

2 In a diving competition, Sally's final dive was awarded the following scores:
8.8 9.1 8.9 9.0 9.2 8.6 8.8

  a Find the mean of the 7 scores given.

 **b** If the highest and lowest scores were left out, what would be the average of the 5 remaining scores?

3 The column graph represents the value of one month's sales at Stan's Super Savings Store.

  a **i** What goods represent the highest value of electrical items sold?

    ii Give two reasons why this may have happened.

  b What was the total value of goods sold?

  c If 200 small kitchen appliances like kettles and toasters were sold, what was their average price?

Electrical items sold
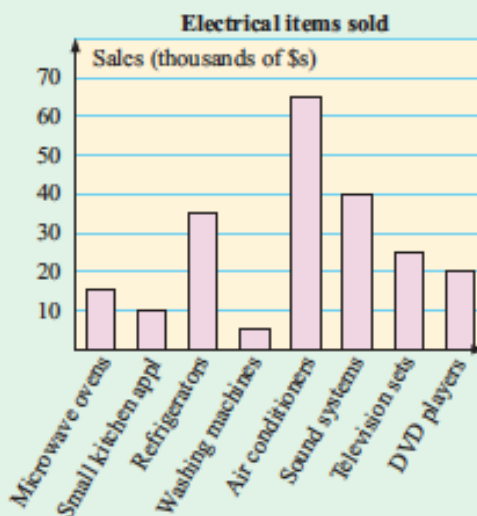
4 Use the pie chart to answer the following questions:

  a What was the major disease causing death as a result of smoking?

  b What 2 groups of diseases made up 50% of all smoking related deaths?

  c If 20 000 people died in one year as a result of smoking, estimate how many died from:

    i heart disease   **ii** lung cancer

    iii other cancers?

Deaths related to smoking
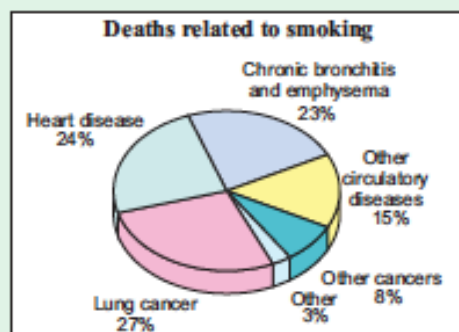
REVIEW SET 17B

1 A medal is awarded to the best and fairest player in a national sporting competition. Umpires award 3 votes to the player they feel was the best and fairest in each game. 2 votes are awarded for second best, and 1 vote for the third best.

Listed below are the votes awarded to a recent winner. The first vote from the left was for the first game, the second vote was for the second game, and so on.

0 2 0 3 1 0 3 2 3 1 1 3 2 0 3 2 1 0 3 2

  a Construct a frequency table showing the votes awarded to the winner in each game.

  b Draw a column graph to show the frequency of the votes.

  c In how many games did the winner not receive votes?

  d What was the winner's total vote?

  e In what percentage of games did the winner receive votes?

  f What was the mean number of votes the winner received per game?

2 The given pie chart represents the sale of €100 000 worth of goods by a sports store during its February sale.

  a Gear for which sport sold best?

  b What value of goods for water sports and tennis was sold?

  c The same percentage of goods was sold in the sports store's $\frac{1}{2}$ million euro 'End of Year Sale'.

    i What value of fitness gear was sold?

    ii What was the total amount of tennis and water sports sales?
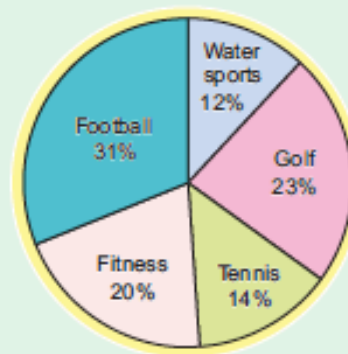
3 The dot plot shows the types of pets treated at a vet on one day.

  a How many pets were treated on this day?

  b Find the mode of the data.

  c What percentage of the pets treated were fish?

Patients at the vet
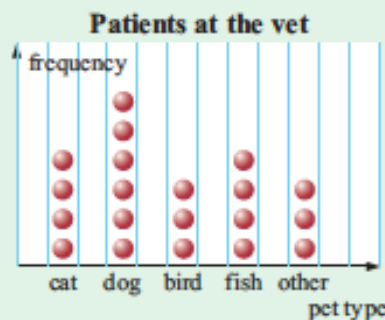
4 Bill's Bakery advertises a new variety in its range of pastries. The daily sales of the new variety are: 23, 25, 18, 21, 17, 14, 15, 19, 18, 11, 15, 12, 6, 9.
Find the mean of the daily sales.

5 The time in minutes taken for customers at a restaurant to receive their meals is given below:

15 28 31 8 22 18 35 24 15 9 28 17 21 20 13

  a Construct a stem-and-leaf plot to display this data.

  b Find the average time for the customers to receive their meals.